THE GENERAL AND LOGICAL THEORY OF AUTOMATA

JOHN VON NEUMANN [*]

The Institute for Advanced Study

I have to ask your forbearance for appearing here, since I am an outsider to most of the fields which form the subject of this conference. Even in the area in which I have some experience, that of the logics and structure of automata, my connections are almost entirely on one side, the mathematical side. The usefulness of what I am going to say, if any, will therefore be limited to this: I may be able to give you a picture of the mathematical approach to these problems, and to prepare you for the experiences that you will have when you come into closer contact with mathematicians. This should orient you as to the ideas and the attitudes which you may then expect to encounter. I hope to get your judgment of the modus procedendi and the distribution of emphases that I am going to use. I feel that I need instruction even in the limiting area between our fields more than you do, and I hope that I shall receive it from your criticisms.

Automata have been playing a continuously increasing, and have by now attained a very considerable, role in the natural sciences. This is a process that has been going on for several decades. During the last part of this period automata have begun to invade certain parts of mathematics too-particularly, but not exclusively, mathematical physics or applied mathematics. Their role in mathematics presents an interesting counterpart to certain functional aspects of organization in nature. Natural organisms are, as a rule, much more complicated and subtle, and therefore much less well understood in detail, than are artificial automata. Nevertheless, some regularities which we observe in the organization of the former may be quite instructive in our thinking and planning of the latter; and conversely, a good deal of our experiences and difficulties with our artificia l automata can be to some extent projected on our interpretations of natural organisms.

PRELIMINARY CONSIDERATIONS

Dichotomy of the Problem: Nature of the Elements, Axiomatic Discussion of Their Synthesis. In comparing living organisms, and, in particular, that most complicated organism, the human central nervous system, with artificial automata, the following limitation should be kept in mind. The natural systems are of enormous complexity, and it is clearly necessary to subdivide the problem that they represent into several parts. One method of subdivision, which is particularly significant in the present context, is this: The organisms can be viewed as made up of parts which to a certain extent are independent, elementary units. We may, therefore, to this extent, view as the first part of the problem the structure and functioning of such elementary units individually. The second part of the problem consists of understanding how these elements are organized into a whole, and how the functioning of the whole is expressed in terms of these elements.

The first part of the problem is at present the dominant one in physiology. It is closely connected with the most difficult chapters of organic chemistry and of physical chemistry, and may in due course be greatly helped by quantum mechanics. I have little qualification to talk about it, and it is not this part with which I shall concern myself here.

The second part, on the other hand, is the one which is likely to attract those of us who have the background and the tastes of a mathematician or a logician. With this attitude, we will be inclined to

[*] This paper is an only slightly edited version of one that was read at the Hixon Symposium on September 20, 1948, in Pasadena, California. Since it was delivered as a single lecture, it was not feasible to go into as much detail on every point as would have been desirable for a final publication. In the present write-up it seemed appropriate to follow the dispositions of the talk; therefore this paper, too, is in many places more sketchy than desirable. It is to be taken only as a general outline of ideas and of tendencies. A detailed account will be published on another occasion.

Note_vgo: “John von Neumann – Collected Works” vol. 5: Design of Computers—Theory of Automata and Numeral Analysis, Q. H. Taub, Ed., Pergamon Press 1963, p. 288-328.

remove the first part of the problem by the process of axiomatization, and concentrate on the second one.

The Axiomatic Procedure. Axiomatizing the behavior of the elements means this: We assume that the elements have certain well-defined, outside, functional characteristics; that is, they are to be treated as "black boxes." They are viewed as automatisms the inner structure of which need not be disclosed, but which are assumed to react to certain unambiguously defined stimuli, by certain unambiguously defined responses.

This being understood, we may then investigate the larger organisms that can be built up from these elements, their structure, their functioning, the connections between the elements, and the general theoretical regularities that may be detectable in the complex syntheses of the organisms in question.

I need not emphasize the limitations of this procedure. Investigations of this type may furnish evidence that the system of axioms used is convenient and, at least in its effects, similar to reality. They are, however, not the ideal method, and possibly not even a very effective method, to determine the validity of the axioms. Such determinations of validity belong primarily to the first part of the problem. Indeed they are essentially covered by the properly physiological (or chemical or physical-chemical) determinations of the nature and properties of the elements.

*The Significant Orders of M*agnit*ude.* In spite of these limitations, however, the "second part" as circumscribed above is important and difficult. With any reasonable definition of what constitutes an- element, the natural organisms are very highly complex aggregations of these elements. The number of cells in the human body is somewhere of the general order of 10^{15} or 10^{16}. The number of neurons in the central nervous system is somewhere of the order of 10^{10}. We have absolutely no past experience with systems of this degree of complexity. All artificial automata made by man have numbers of parts which by any comparably schematic count are of the order 10^3 to 10^6. In addition, those artificial systems which function with that type of logical flexibility and autonomy that we find in the natural organisms do not lie at the peak of this scale. The prototypes for these systems are the modem computing machines, and here a reasonable definition of what constitutes an element will lead to counts of a few times 10^3 or 10^4 elements.

DISCUSSION OF CERTAIN RELEVANT TRAITS OF COMPUTING MACHINES

Computing Machines-Typical Operations. Having made these general remarks, let me now be more definite, and turn to that part of the subject about which I shall talk in specific and technical detail. As I have indicated, it is concerned with artificial automata and more specially with computing machines. They have some similarity to the central nervous system, or at least to a certain segment of the system's functions. They are of course vastly less complicated, that is, smaller in the sense which really matters. It is nevertheless of a certain interest to analyze the problem of organisms and organization from the point of view of these relatively small, artificial automata, and to effect their comparisons with the central nervous system from this frog's-view perspective.

I shall begin by some statements about computing machines as such. The notion of using an automaton for the purpose of computing is relatively new. While computing automata are not the most complicated artificial automata from the point of view of the end results they achieve, they do nevertheless represent the highest degree of complexity in the sense that they produce the longest chains of events determining and following each other.

There exists at the present time a reasonably well-defined set of ideas about when it is reasonable to use a fast computing machine, and when it is not. The criterion is usually expressed in terms of the multiplications involved in the mathematical problem. The use of a fast computing machine is believed to be by and large justified when the computing task involves about a million multiplications or more in a sequence.

An expression in more fundamentally logical terms is this: In the relevant fields (that is, in those parts of [usually applied] mathematics, where the use of such machines is proper) mathematical experience indicates the desirability of precisions of about ten decimal places. A single multiplication would therefore seem to involve at least 10 x 10 steps (digital multiplications); hence a million multiplications amount to at least 10^8 operations. Actually, however, multiplying two decimal digits is not an elementary operation. There are various ways of breaking it down into such, and all of them have about the same degree of complexity. The simplest way to estimate this degree of complexity is, instead of counting decimal places, to count the number of places that would be required for the same precision in the binary system of notation (base 2 instead of base 10). A decimal digit corresponds to about three binary digits, hence ten decimals to about thirty binary. The multiplication referred to above, therefore, consists not of 10 x 10, but of 30 x 30 elementary steps, that is, not 10^2, but 10^3 steps. (Binary digits are "all or none" affairs, capable of the values 0 and 1 only. Their multiplication is, therefore, indeed an elementary operation. By the way, the equivalent of 10 decimals is 33 [rather than 30] binaries but 33 x 33, too, is approximately 10^3.) It follows, therefore, that a million multiplications in the sense indicated above are more reasonably described as corresponding to 10^9 elementary operations.

Precision and Reliability Requirements. I am not aware of any other field of human effort where the result really depends on a sequence of a billion (10^9) steps in any artifact, and where, furthermore, it has the characteristic that every step actually matters—or, at least, may matter with a considerable probability. Yet, precisely this is true for computing machines—this is their most specific and most difficult characteristic.

Indeed, there have been in the last two decades automata which did perform hundreds of millions, or even billions, of steps before they produced a result. However, the operation of these automata is not serial. The large number of steps is due to the fact that, for a variety of reasons, it is desirable to do the same experiment over and over again. Such cumulative, repetitive procedures may, for instance, increase the size of the result, that is (and this is the important consideration), increase the significant result, the "signal," relative to the "noise" which contaminates it. Thus any reasonable count of the number of reactions which a microphone gives before a verbally interpretable acoustic signal is produced is in the high tens of thousands. Similar estimates in television will give tens of millions, and in radar possibly many billions. If, however, any of these automata makes mistakes, the mistakes usually matter only to the extent of the fraction of the total number of steps which they represent. (This is not exactly true in all relevant examples, but it represents the qualitative situation better than the opposite statement.) Thus the larger the number of operations required to produce a result, the smaller will be the significant contribution of every individual operation.

In a computing machine no such rule holds. Any step is (or may potentially be) as important as the whole result; any error can vitiate the result in its entirety. (This statement is not absolutely true, but probably nearly 30 per cent of all steps made are usually of this sort.) Thus a computing machine is one of the exceptional artifacts. They not only have to perform a billion or more steps in a short time, but in a considerable part of the procedure (and this is a part that is rigorously specified in advance) they are permitted not a single error. In fact, in order to be sure that the whole machine is operative, and that no potentially degenerative malfunctions have set in, the present practice usually requires that no error should occur anywhere in the entire procedure.

This requirement puts the large, high-complexity computing machines in an altogether new light. It makes in particular a comparison between the computing machines and the operation of the natural organisms not entirely out of proportion.

The Analogy Principle. All computing automata fall into two great classes in a way which is immediately obvious and which, as you will see in a moment, carries over to living organisms. This classification is into analogy and digital machines.

Let us consider the analogy principle first. A computing machine may be based on the principle that numbers are represented by certain physical quantities. As such quantities we might, for instance, use the intensity of an electrical current, or the size of an electrical potential, or the number of degrees of arc by which a disk has been rotated (possibly in conjunction with the number of entire revolutions effected), etc. Operations like addition, multiplication, and integration may then be performed by finding various natural processes which act on these quantities in the desired way. Currents may be multiplied by feeding them into the two magnets of a dynamometer, thus producing a rotation. This rotation may then be transformed into an electrical resistance by the attachment of a rheostat; and, finally, the resistance can be transformed into a current by connecting it to two sources of fixed (and different) electrical potentials. The entire aggregate is thus a "black box" into which two currents are fed and which produces a current equal to their product. You are certainly familiar with many other ways in which a wide variety of natural processes can be used to perform this and many other mathematical operations.

The first well-integrated, large computing machine ever made was an analogy machine, V. Bush's Differential Analyzer. This machine, by the way, did the computing not with electrical currents, but with rotating disks. I shall not discuss the ingenious tricks by which the angles of rotation of these disks were combined according to various operations of mathematics.

I shall make no attempt to enumerate, classify, or systematize the wide variety of analogy principles and mechanisms that can be used in computing. They are confusingly multiple. The guiding principle without which it is impossible to reach an understanding of the situation is the classical one of all "communication theory"—the "signal to noise ratio." That is, the critical question with every analogy procedure is this: How large are the uncontrollable fluctuations of the mechanism that constitute the "noise," compared to the significant "signals" that express the numbers on which the machine operates? The usefulness of any analogy principle depends on how low it can keep the relative size of the uncontrollable fluctuations—the "noise level."

To put this in another way. No analogy machine exists which will really form the product of two numbers. What it will form is this product, plus a small but unknown quantity which represents the random noise of the mechanism and the physical processes involved. The whole problem is to keep this quantity down. This principle has controlled the entire relevant technology. It has, for instance, caused the adoption of seemingly complicated and clumsy mechanical devices instead of the simpler and elegant electrical ones. (This, at least, has been the case throughout most of the last twenty years. More recently, in certain applications which required only very limited precision the electrical devices have again come to the fore.) In comparing mechanical with electrical analogy processes, this roughly is true: Mechanical arrangements may bring this noise level below the "maximum signal level" by a factor of something like $1:10^4$ or 10^5. In electrical arrangements, the ratio is rarely much better than $1:10^2$. These ratios represent, of course, errors in the elementary steps of the calculation, and not in its final results. The latter will clearly be substantially larger.

The Digital Principle. A digital machine works with the familiar method of representing numbers as aggregates of digits. This is, by the way, the procedure which all of us use in our individual, non-mechanical computing, where we express numbers in the decimal system. Strictly speaking, digital computing need not be decimal. Any integer larger than one may be used as the basis of a digital notation for numbers. The decimal system (base 10) is the most common one, and all digital machines built to date operate in this system. It seems likely, however, that the binary (base 2) system will, in the end, prove preferable, and a number of digital machines using that system are now under construction.

The basic operations in a digital machine are usually the four species of arithmetic: addition, subtraction, multiplication, and division. We might at first think that, in using these, a digital machine possesses (in contrast to the analogy machines referred to above) absolute precision. This, however, is not the case, as the following consideration shows.

Take the case of multiplication. A digital machine multiplying two 10-digit numbers will produce a 20-digit number, which is their product, with no error whatever. To this extent its precision is absolute, even though the electrical or mechanical components of the arithmetical organ of the machine are as such of limited precision. As long as there is no breakdown of some component, that is, as long as the operation of each component produces only fluctuations within its preassigned tolerance limits, the result will be absolutely correct. This is, of course, the great and characteristic virtue of the digital procedure. Error, as a matter of normal operation and not solely (as indicated above) as an accident attributable to some definite breakdown, nevertheless creeps in, in the following manner. The absolutely correct product of two 10-digit numbers is a 20-digit number. If the machine is built to handle 10-digit numbers only, it will have to disregard the last 10 digits of this 20-digit number and work with the first 10 digits alone. (The small, though highly practical, improvement due to a possible modification of these digits by "round-off" may be disregarded here.) If, on the other hand, the machine can handle 20-digit numbers, then the multiplication of two such will produce 40 digits, and these again have to be cut down to 20, etc., etc. (To conclude, no matter what the maximum number of digits is for which the machine has been built, in the course of successive multiplications this maximum will be reached, sooner or later. Once it has been reached, the next multiplication will produce supernumerary digits, and the product will have to be cut to half of its digits [the first half, suitably rounded off]. The situation for a maximum of 10 digits is therefore typical, and we might as well use it to exemplify things.)

Thus the necessity of rounding off an (exact) 20-digit product to the regulation (maximum) number of 10 digits introduces in a digital machine qualitatively the same situation as was found above in an analogy machine. What it produces when a product is called for is not that product itself, but rather the product plus a small extra term—the round-off error. This error is, of course, not a random variable like the noise in an analogy machine. It is, arithmetically, completely determined in every particular instance. Yet its mode of determination is so complicated, and its variations throughout the number of instances of its occurrence in a problem so irregular, that it usually can be treated to a high degree of approximation as a random variable.

(These considerations apply to multiplication. For division the situation is even slightly worse, since a quotient can, in general, not be expressed with absolute precision by any finite number of digits. Hence here rounding off is usually already a necessity after the first operation. For addition and subtraction, on the other hand, this difficulty does not arise: The sum or difference has the same number of digits if there is no increase in size beyond the planned maximum] as the addends themselves. Size may create difficulties which are added to the difficulties of precision discussed here, but I shall not go into these at this time.)

The Role of the Digital Procedure in Reducing the Noise Level. The important difference between the noise level of a digital machine, as described above, and of an analogy machine is not qualitative at all; it is quantitative. As pointed out above, the relative noise level of an analogy machine is never lower than 1 in 10^3, and in many cases as high as 1 in 10^2. In the 10-place decimal digital machine referred to above the relative noise level (due to round-off) is 1 part in 10^2. Thus the real importance of the digital procedure lies in its ability to reduce the computational noise level to an extent which is completely unobtainable by any other (analogy) procedure. In addition, further reduction of the noise level is increasingly difficult in an analogy mechanism, and increasingly easy in a digital one. In all analogy machine a precision of 1 in 10^3 is easy to achieve; 1 in 10^4 somewhat difficult; 1 in 10^5 very difficult; and 1 in 10^6 impossible, in the present state of technology. In a digital machine, the above precisions mean merely that one builds the machine to 3, 4, 5, and 6 decimal places, respectively. Here the transition from each stage to the next one gets actually easier. Increasing a 3-place machine (if anyone wished to build such a machine) to a 4-place machine is a 33 per cent increase; going from 4 to 5 places, a 20 per cent increase; going from 5 to 6 places, a 17 per cent increase. Going from 10 to 11 places is only a 10 per cent increase. This is clearly an entirely different milieu, from the point of view of the reduction of "random noise," from that of physical processes. It is here—and not in its practically ineffective absolute reliability—that the importance of the digital procedure lies.

COMPARISONS BETWEEN COMPUTING MACHINES AND LIVING ORGANISMS

Mixed Character of Living Organisms. When the central nervous system is examined, elements of both procedures; digital and analogy, are discernible.

The neuron transmits an impulse. This appears to be its primary function, even if the last word about this function and its exclusive or non-exclusive character is far from having been said. The nerve impulse seems in the main to be an all-or-none affair, comparable to a binary digit. Thus a digital element is evidently present, but it is equally evident that this is not the entire story. A great deal of what goes on in the organism is not mediated in this manner, but is dependent on the general chemical composition of the blood stream or of other humoral media. It is well known that there are various composite functional sequences in the organism which have to go through a variety of steps from the original stimulus to the ultimate effect—some of the steps being neural, that is, digital, and others humoral, that is, analogy. These digital and analogy portions in such a chain may alternately multiply. In certain cases of this type, the chain can actually feed back into itself, that is, its ultimate output may again stimulate its original input.

It is well known that such mixed (part neural and part humoral) feedback chains call produce processes of great importance. Thus the mechanism which keeps the blood pressure constant is of this mixed type. The nerve which senses and reports the blood pressure does it by a sequence of neural impulses, that is, in a digital manner. The muscular contraction which this impulse system induces may still be described as a superposition of many digital impulses. The influence of such a contraction on the blood stream is, however, hydrodynamical, and hence analogy. The reaction of the pressure thus produced back on the nerve which reports the pressure closes the circular feedback, and at this point the analogy procedure again goes over into a digital one. The comparisons between the living organisms and the computing machines are, therefore, certainly imperfect at this point. The living organisms are very complex—part digital and part analogy—mechanisms. The computing machines, at least in their recent forms to which I am referring in this discussion, are purely digital. Thus I must ask you to accept this oversimplification of the system. Although I am well aware of the analogy component in living organisms, and it would be absurd to deny its importance, I shall, nevertheless, for the sake of the simpler discussion, disregard that part. I shall consider the living organisms as if they were purely digital automata.

Mixed Character of Each Element. In addition to this, one may argue that even the neuron is not exactly a digital organ. This point has been put forward repeatedly and with great force. There is certainly a great deal of truth in it, when one considers things in considerable detail. The relevant assertion is, in this respect, that the fully developed nervous impulse, to which all-or-none character can be attributed, is not an elementary phenomenon, but is highly complex. It is a degenerate state of the complicated electrochemical complex which constitutes the neuron, and which in its fully analyzed functioning must be viewed as an analogy machine. Indeed, it is possible to stimulate the neuron in such a way that the breakdown that releases the nervous stimulus will not occur. In this area of “subliminal stimulation”, we find first (that is, for the weakest stimulations) responses which are proportional to the stimulus, and then (at higher, but still subliminal, levels of stimulation) responses which depend on more complicated non-linear laws, but are nevertheless continuously variable and not of the breakdown type. There are also other complex phenomena within and without the subliminal range: fatigue, summation, certain forms of self-oscillation, etc.

In spite of the truth of these observations, it should be remembered that they may represent an improperly rigid critique of the concept of an all-or-none organ. The electromechanical relay, or the vacuum tube, when properly used, are undoubtedly all-or-none organs. Indeed, they are the prototypes of such organs. Yet both of them are in reality complicated analogy mechanisms, which upon appropriately adjusted stimulation respond continuously, linearly or non-linearly, and exhibit the phenomena of "breakdown" or "all-or-none-response only under very particular conditions of operation. There is little difference between this performance and the above-described performance of neurons. To put it somewhat differently. None of these is an exclusively all-or-none organ (there is little in our technological or physiological experience to indicate that absolute all-or-none organs exist); this, however, is irrelevant. By an all-or-none organ we should rather mean one which fulfills the following two conditions. First, it functions in the all-or-none manner under certain suitable operating conditions. Second, these operating conditions are the ones under which it is normally used; they represent the functionally normal state of affairs within the large organism, of which it forms a part. Thus the important fact is not whether an organ has necessarily and under all conditions the all-or-none character—this is probably never the case—but rather whether in its proper context it functions primarily, and appears to be intended to function primarily, as an all-or-none organ. I realize that this definition brings in rather undesirable criteria of "propriety" of context, of "appearance" and "intention." I do not see, however, how we can avoid using them, and how we can forego counting on the employment of common sense in their application. I shall, accordingly, in what follows use the working hypothesis that the neuron is an all-or-none digital organ. I realize that the last word about this has not been said, but I hope that the above excursus on the limitations of this working hypothesis and the reasons for its use will reassure you. I merely want to simplify my discussion; I am not trying to prejudge any essential open question.

In the same sense, I think that it is permissible to discuss the neurons as electrical organs. The stimulation of a neuron, the development and progress of its impulse, and the stimulating effects of the impulse at a synapse can all be described electrically. The concomitant chemical and other processes are important in order to understand the internal functioning of a nerve cell. They may even be more important than the electrical phenomena. They seem, however, to be hardly necessary for a description of a neuron as a "black box," an organ of the all-or-none type. Again the situation is no worse here than it is for, say, a vacuum tube. Here, too, the purely electrical phenomena are accompanied by numerous other phenomena of solid state physics, thermodynamics, mechanics. All of these are important to understand the structure of a vacuum tube, but are best excluded from the discussion, if it is to treat the vacuum tube as a "black box" with a schematic description.

The Concept of a Switching Organ or Relay Organ. The neuron, as well as the vacuum tube, viewed under the aspects discussed above, are then two instances of the same generic entity, which it is customary to call a "switching organ" or "relay organ." (The electromechanical relay is, of course, another instance.) Such an organ is defined as a "black box," which responds to a specified stimulus or combination of stimuli by an energetically independent response. That is, the response is expected to have enough energy to cause several stimuli of the same kind as the ones which initiated it. The energy of the response, therefore, cannot have been supplied by the original stimulus. It must originate in a different and independent source of power. The stimulus merely directs, controls the flow of energy from this source.

(This source, in the case of the neuron, is the general metabolism of the neuron. In the case of a vacuum tube, it is the power which maintains the cathode-plate potential difference, irrespective of whether the tube is conducting or not, and to a lesser extent the heater power which keeps "boiling" electrons out of the cathode. In the case of the electromechanical relay, it is the current supply whose path the relay is closing or opening.)

The basic switching organs of the living organisms, at least to the extent to which we are considering them here, are the neurons. The basic switching organs of the recent types of computing machines are vacuum tubes; in older ones they were wholly or partially electromechanical relays. It is quite possible that computing machines will not always be primarily aggregates of switching organs, but such a development is as yet quite far in the future. A development which may lie much closer is that the vacuum tubes may be displaced from their role of switching organs in computing machines. This, too, however, will probably not take place for a few years yet. I shall, therefore, discuss computing machines solely from the point of view of aggregates of switching organs which are vacuum tubes.

Comparison of the Sizes of Large Computing Machines and Living Organisms. Two well-known, very large vacuum tube computing machines are in existence and in operation. Both consist of about 20,000 switching organs. One is a pure vacuum tube machine. (It belongs to the U. S. Army Ordnance Department, Ballistic Research Laboratories, Aberdeen, Maryland, designation "ENIAC.") The other is mixed-part vacuum tube and part electromechanical relays. (It belongs to the I. B. M. Corporation, and is located in New York, designation "SSEC.") These machines are a good deal larger than what is likely to be the size of the vacuum tube computing machines which will come into existence and operation in the next few years. It is probable that each one of these will consist of 2000 to 6000 switching organs. (The reason for this decrease lies in a different attitude about the treatment of the "memory," which I will not discuss here.) It is possible that in later years the machine sizes will increase again, but it is not likely that 10,000 (or perhaps a few times 10,000) switching organs will be exceeded as long as the present techniques and philosophy are employed. To sum up, about 10^4 switching organs seem to be the proper order of magnitude for a computing machine.

In contrast to this, the number of neurons in the central nervous system has been variously estimated as something of the order of 10^{10} do not know how good this figure is, but presumably the exponent at least is not too high, and not too low by more than a unit. Thus it is very conspicuous that the central nervous system is at least a million times larger than the largest artificial automaton that we can talk about at present. It is quite interesting to inquire why this should be so and what questions of principle are involved. It seems to me that a few very clear-cut questions of principle are indeed involved.

Determination of the Significant Ratio of Sizes fur the Elements.

Obviously, the vacuum tube, as we know it, is gigantic compared to a nerve cell. Its physical volume is about a billion times larger, and its energy dissipation is about a billion times greater. (It is, of course, impossible to give such figures with a unique validity, but the above ones are typical.) There is, on the other hand, a certain compensation for this. Vacuum tubes can be made to operate at exceedingly high speeds in applications other than computing machines, but these need not concern us here. In computing machines the maximum is a good deal lower, but it is still quite respectable. In the present state of the art, it is generally believed to be somewhere around a million actuations per second. The responses of a nerve cell are a good deal slower than this, perhaps 1/2000 of a second, and what really matters, the minimum time-interval required from stimulation to complete recovery and, possibly, renewed stimulation, is still longer than this—at best approximately 1/200 of a second. This gives a ratio of 1:5000, which, however, may he somewhat too favorable to the vacuum tube, since vacuum tubes, when used as switching organs at the 1,000,000 steps per second rate, are practically never run at a 100 per cent duty cycle. A ratio like 1:2000 would, therefore, seem to be more equitable. Thus the vacuum tube, at something like a billion times the expense, outperforms the neuron by a factor of somewhat over 1000. There is, therefore, some justice in saying that it is less efficient by a factor of the order of a million.

The basic fact is, in every respect, the small size of the neuron compared to the vacuum tube. This ratio is about a billion, as pointed out above. What is it due to?

Analysis of the Reasons for the Extreme Ratio of Sizes. The origin of this discrepancy lies in the fundamental control organ or, rather, control arrangement of the vacuum tube as compared to that of the neuron. In the vacuum tube the critical area of control is the space between the cathode (where the active agents, the electrons, originate) and the grid (which controls the electron flow). This space is about one millimeter deep. The corresponding entity in a neuron is the wall of the nerve cell, the "membrane." Its thickness is about a micron (1/1000 millimeter), or somewhat less. At this point, therefore, there is a ratio of approximately 1:1000 in linear dimensions. This, by the way, is the main difference. The electrical fields, which exist in the controlling space, are about the same for the vacuum tube and for the neuron. The potential differences by which these organs can be reliably steered are tens of volts in one case and tens of millivolts in the other. Their ratio is again about 1:1000, and hence their gradients (the field strengths) are about identical. Now a ratio of 1:1000 in linear dimensions corresponds to a ratio of 1:1,000,000,000 in volume. Thus the discrepancy factor of a billion in 3-dimensional size (volume) corresponds, as it should, to a discrepancy factor of 1000 in linear size, that is, to the difference between the millimeter inter-electrode-space depth of the vacuum tube and the micron membrane thickness of the neuron.

It is worth noting, although it is by no means surprising, how this divergence between objects, both of which are microscopic and are situated in the interior of the elementary components leads to impressive macroscopic differences between the organisms built upon them. This difference between a millimeter object and a micron object causes the ENIAC to weigh 30 tons and to dissipate 150 kilowatts of energy, while the human central nervous system, which is functionally about a million times larger, has the weight of the order of a pound and is accommodated within the human skull. In assessing the weight and size of the ENIAC as stated above, we should also remember that this huge apparatus is needed in order to handle 20 numbers of 10 decimals each, that is, a total of 200 decimal digits, the equivalent of about 700 binary digits—merely 700 simultaneous pieces of "yes-no" information)

Technological Interpretation of *These Reasons.* These considerations should make it clear that our present technology is still very imperfect in handling information at high speed .and high degrees of complexity. The apparatus which results is simply enormous, both physically and in its energy requirements.

The weakness of this technology lies probably, in part at least, in the materials employed. Our present techniques involve the using of metals, with rather close spacings, and at certain critical points separated by vacuum only. This combination of media has a peculiar mechanical instability that is entirely alien to living nature. By this I mean the simple fact that, if a living organism is mechanically injured, it has a strong tendency to restore itself. If, on the other hand, we hit a man-made mechanism with a sledge hammer, no such restoring tendency is apparent. If two pieces of metal are close together, the small vibrations and other mechanical disturbances, which always exist in the ambient medium, constitute a risk in that they may bring them into contact. If they were at different electrical potentials, the next thing that may happen after this short circuit is that they can become electrically soldered together and the contact becomes permanent. At this point, then, a genuine and permanent breakdown will have occurred. When we injure the membrane of a nerve cell, no such thing happens. On the contrary, the membrane will usually reconstitute itself after a short delay.

It is this mechanical instability of our materials which prevents us from reducing sizes further. This instability and other phenomena of a comparable character make the behavior in our componentry less than wholly reliable, even at the present sizes. Thus it is the inferiority of our materials, compared with those used in nature, which prevents us from attaining the high degree of complication and the small dimensions which have been attained by natural organisms.

THE FUTURE LOGICAL THEORY OF AUTOMATA

Further Discussion of the Factors That Limit the Present Size of Artificial Automata. We have emphasized how the complication is limited in artificial automata, that is, the complication which can be handled without extreme difficulties and for which automata can still be expected to function reliably. Two reasons that put a limit on complication in this sense have already been given. They are the large size and the limited reliability of the componentry that we must use, both of them due to the fact that we are employing materials which seem to be quite satisfactory in simpler applications, but. marginal and inferior to the natural ones in this highly complex application. There is, however, a third important limiting factor, and we should now turn our attention to it. This factor is of an intellectual, and not physical, character.

The Limitation Which Is Due to the Lack of a Logical Theory of Automata. We are very far from possessing a theory of automata which deserves that name, that is, a properly mathematical-logical theory. There exists today a very elaborate system of formal logic, and, specifically, of logic as applied to mathematics. This is a discipline with many good sides, but also with certain serious weaknesses. This is not the occasion to enlarge upon the good sides, which I have certainly no intention to belittle. About the inadequacies, however, this may be said: Everybody who has worked in formal logic will confirm that it is one of the technically most refractory parts of mathematics. The reason for this is that it deals with rigid, all-or-none concepts, and has very little contact with the continuous concept of the real or of the complex number, that is, with mathematical analysis. Yet analysis is the technically most successful and best-elaborated part of mathematics. Thus formal logic is, by the nature of its approach, cut off from the best cultivated portions of mathematics, and forced onto the most difficult part of the mathematical terrain, into combinatorics.

The theory of automata, of the digital, all-or-none type, as discussed up to now, is certainly a chapter in formal logic. It would, therefore, seem that it will have to share this unattractive property of formal logic. It will have to be, from the mathematical point of view, combinatorial rather than analytical.

Probable Characteristics of Such a Theory. Now it seems to me that this will in fact not be the case. In studying the functioning of automata, it is clearly necessary to pay attention to a circumstance which has never before made its appearance in formal logic.

Throughout all modern logic, the only thing that is important is whether a result can be achieved in a finite number of elementary steps or not. The size of the number of steps which are required, on the other hand, is hardly ever a concern of formal logic. Any finite sequence of correct steps is, as a matter of principle, as good as any other. 1t is a matter of no consequence whether the number is small or large, or even so large that it couldn't possibly be carried out in a lifetime, or in the presumptive lifetime of the stellar universe as we know it. In dealing with automata, this statement must he significantly modified. In the case of an automaton the thing which matters is not only whether it can reach a certain result in a finite number of steps at all but also how many such steps are needed. There are two reasons. First, automata are constructed in order to reach certain results in certain pre-assigned durations, or at least in pre-assigned orders of magnitude of duration. Second, the componentry employed has in every individual operation a small but nevertheless non-zero probability of failing. In a sufficiently long chain of operations the cumulative effect of these individual probabilities of failure may (if unchecked) reach the order of magnitude of unity—at which point it produces, in effect, complete unreliability. The probability levels which are involved here are very low, but still not too far removed from the domain of ordinary technological experience. It is not difficult to estimate that a high-speed computing machine, dealing with a typical problem, may have to perform as much as 10^{12} individual operations. The probability of error on an individual operation which can be tolerated must, therefore, be small compared to 10^{-12}. I might mention that an electromechanical relay (a telephone relay) is at present considered acceptable if its probability of failure on an individual operation is of the order 10^{-8}. It is considered excellent if this order of probability is 10^{-9}. Thus the reliabilities required in a high-speed computing machine are higher, but not prohibitively higher, than those that constitute sound practice in certain existing industrial fields. The actually obtainable reliabilities are, however, not likely to leave a very wide margin against the minimum requirements just mentioned. An exhaustive study and a nontrivial theory will, therefore, certainly be called for.

Thus the logic of automata will differ from the present system of formal logic in two relevant respects.

1. The actual length of "chains of reasoning," that is, of the chains of operations, will have to be considered.

2. The operations of logic (syllogisms, conjunctions, disjunctions, negations, etc., that is, in the terminology that is customary for automata, various forms of gating, coincidence, anti-coincidence, blocking, etc., actions) will all have to be treated by procedures which allow exceptions (malfunctions) with low but non-zero probabilities. All of this will lead to theories which are much less rigidly of an all-or-none nature than past and present formal logic. They will be of a much less combinatorial, and much more analytical, character. In fact, there are numerous indications to make us believe that this new system of formal logic will move closer to another discipline which has been little linked in the past with logic. This is thermodynamics, primarily in the form it was received from Boltzmann, and is that part of theoretical physics which comes nearest in some of its aspects to manipulating and measuring information. Its techniques are indeed much more analytical than combinatorial, which again illustrates the point that I have been trying to make above. It would, however, take me too far to go into this subject more thoroughly on this occasion.

All of this re-emphasizes the conclusion that was indicated earlier, that a detailed, highly mathematical, and more specifically analytical, theory of automata and of information is needed. We possess only the first indications of such a theory at present. In assessing artificial automata, which are, as I discussed earlier, of only moderate size, it has been possible to get along in a rough, empirical manner without such a theory. There is every reason to believe that this will not be possible with more elaborate automata.

Effects of the Lack of a Logical Theory of Automata on the Procedures in Dealing with Errors. This, then, is the last, and very important, limiting factor. It is unlikely that we could construct automata of a much higher complexity than the ones we now have, without possessing a very advanced and subtle theory of automata and information. A fortiori, this is inconceivable for automata of such enormous complexity as is possessed by the human central nervous system.

This intellectual inadequacy certainly prevents us from getting much farther than we are now.

A simple manifestation of this factor is our present relation to error checking. In living organisms malfunctions of components occur. The organism obviously has a way to detect them and render them harmless. It is easy to estimate that the number of nerve actuations which occur in a normal lifetime must be of the order of 10^{20}. Obviously, during this chain of events there never occurs a malfunction which cannot be corrected by the organism itself, without any significant outside intervention. The system must, therefore, contain the necessary arrangements to diagnose errors as they occur, to readjust the organism so as to minimize the effects of the errors, and finally to correct or to block permanently the faulty components. Our modus procedendi with respect to malfunctions in our artificial automata is entirely different. Here the actual practice, which has the consensus of all experts of the field, is somewhat like this: Every effort is made to detect (by mathematical or by automatical checks) every error as soon as it occurs. Then an attempt is made to isolate the component that caused the error as rapidly as feasible. This may be done partly automatically, but in any case a significant part of this diagnosis must be effected by intervention from the outside. Once the faulty component has been identified, it is immediately corrected or replaced.

Note the difference in these two attitudes. The basic principle of dealing with malfunctions in nature is to make their effect as unimportant as possible and to apply correctives, if they are necessary at all, at leisure. 1n our dealings with artificial automata, on the other hand, we require an immediate diagnosis. Therefore, we are trying to arrange the automata in such a manner that errors will become as conspicuous as possible, and intervention and correction follow immediately. In other words, natural organisms are constructed to make errors as inconspicuous, as harmless, as possible. Artificial automata are designed to make errors as conspicuous, as disastrous, as possible. The rationale of this difference is not far to seek. Natural organisms are sufficiently well conceived to be able to operate even when malfunctions have set in. They can operate in spite of malfunctions, and their subsequent tendency is to remove these malfunctions. An artificial automaton could certainly be designed so as to be able to operate normally in spite of a limited number of malfunctions in certain limited areas. Any malfunction, however, represents a considerable risk that some generally degenerating process has already set in within the machine. It is, therefore, necessary to intervene immediately, because a machine which has begun to malfunction has only rarely a tendency to restore itself, and will more probably go from bad to worse. All of this comes back to one thing. With our artificial automata we are moving much more in the dark than nature appears to be with its organisms. We are, and apparently, at least at present, have to be, much more "scared" by the occurrence of an isolated error and by the malfunction which must be behind it. Our behavior is clearly that of overcaution, generated by ignorance.

The Single-Error Principle. A minor side light to this is that almost all our error-diagnosing techniques are based on the assumption that the machine contains only one faulty component. In this case, iterative subdivisions of the machine into parts permit us to determine which portion contains the fault. As soon as the possibility exists that the machine may contain several faults, these, rather powerful, dichotomic methods of diagnosis are lost. Error diagnosing then becomes an increasingly hopeless proposition. The high premium on keeping the number of errors to be diagnosed down to one, or at any rate as low as possible, again illustrates our ignorance in this field, and is one of the main reasons why errors must be made as conspicuous as possible, in order to be recognized and apprehended as soon after their occurrence as feasible, that is, before further errors have had time to develop.

PRINCIPLES OF DIGITALIZATION

Digitalization of Continuous Quantities: the Digital Expansion Method and the Counting Method. Consider the digital part of a natural organism; specifically, consider the nervous system. It seems that we are indeed justified in assuming that this is a digital mechanism, that it transmits messages which are made up of signals possessing the all-or-none character. (See also the earlier discussion, page 10.) In other words, each elementary signal, each impulse, simply either is or is not there, with no further shadings. A particularly relevant illustration of this fact is furnished by those cases where the underlying problem has the opposite character, that is, where the nervous system is actually called upon to transmit a continuous quantity. Thus the case of a nerve which has to report on the value of a pressure is characteristic.

Assume, for example, that a pressure (clearly a continuous quantity) is to be transmitted. It is well known how this trick is done. The nerve which does it still transmits nothing but individual all-or-none impulses. How does it then express the continuously numerical value of pressure in terms of these impulses, that is, of digits? In other words, how does it encode a continuous number into a digital notation? It does certainly not do it by expanding the number in question into decimal (or binary, or any other base) digits in the conventional sense. What appears to happen is that it transmits pulses at a frequency which varies and which is within certain limits proportional to the continuous quantity in question, and generally a monotone function of it. The mechanism which achieves this "encoding" is, therefore, essentially a frequency modulation system.

The details are known. The nerve has a finite recovery time. In other words, after it has been pulsed once, the time that has to lapse before another stimulation is possible is finite and dependent upon the strength of the ensuing (attempted) stimulation. Thus, if the nerve is under the influence of a continuing stimulus (one which is uniformly present at all times, like the pressure that is being considered here), then the nerve will respond periodically, and the length of the period between two successive stimulations is the recovery time referred to earlier, that is, a function of the strength of the constant stimulus (the pressure in the present case). Thus, under a high pressure, the nerve may be able to respond every 8 milliseconds, that is, transmit at the rate of 125 impulses per second; while under the influence of a smaller pressure it may be able to repeat only every 14 milliseconds, that is, transmit at the rate of 71 times per second. This is very clearly the behavior of a genuinely yes-or-no organ, of a digital organ. It is very instructive, however, that it uses a "count" rather than a "decimal expansion" (or "binary expansion," etc.) method.

Comparison of the Two Methods. The Preference of Living Organisms for the Counting Method. Compare the merits and demerits of these two methods. The counting method is certainly loss efficient than the expansion method. In order to express a number of about a million (that is, a physical quantity of a million distinguishable resolution steps) by counting, a million pulses have to be transmitted. In order to express a number of the same size by expansion, 6 or 7 decimal digits are needed,. that is, about 20 binary digits. Hence, in this case only 20 pulses are needed. Thus our expansion method is much more economical in notation than the counting methods which are resorted to by nature. On the other hand, the counting method has a high stability and safety from error. If you express a number of the order of a million by counting and miss a count, the result is only irrelevantly changed. If you express it by (decimal or binary) expansion, a single error in a single digit may vitiate the entire result. Thus the undesirable trait of our computing machines reappears in our digital expansion system; in fact, the former is clearly deeply connected with, and partly a consequence of, the latter. The high stability and nearly error-proof character of natural organisms, on the other hand, is reflected in the counting method that they seem to use in this case. All of this reflects a general rule. You can increase the safety from error by a reduction of the efficiency of the notation, or, to say it positively, by allowing redundancy of notation. Obviously, the simplest form of achieving safety by redundancy is to use the, per se, quite unsafe digital expansion notation, but to repeat every such message several times. In the case under discussion, nature has obviously resorted to an even more redundant and even safer system.

There are, of course, probably other reasons why the nervous system uses the counting rather than the digital expansion. The encoding-decoding facilities required by the former are much simpler than those required by the latter. It is true, however, that nature seems to be willing and able to go much further in the direction of complication than we are, or rather than we can afford to go. One may, therefore, suspect that if the only demerit of the digital expansion system were its greater logical complexity, nature would not, for this reason alone, have rejected it. It is, nevertheless, trite that we have nowhere an indication of its use in natural organisms. It is difficult to tell how much "final" validity one should attach to this observation. The point deserves at any rate attention, and should receive it in future investigations of the functioning of the nervous system.

FORMAL NEURAL NETWORKS

The McCulloch-Pitts Theory of Formal Neural Networks. A great deal more could be said about these things from the logical and the organizational point of view, but I shall not attempt to say it here. I shall instead go on to discuss what is probably the most significant result obtained with the axiomatic method up to now. I mean the remarkable theorems of McCulloch and Pitts on the relationship of logics and neural networks.

In this discussion I shall, as I have said, take the strictly axiomatic point of view. I shall, therefore, view a neuron as a "black box" with a certain number of inputs that receive stimuli and an output that emits stimuli. To be specific, I shall assume that the input connections of each one of these can be of two types, excitatory and inhibitory. The boxes themselves are also of two types, threshold 1 and threshold 2. These concepts are linked and circumscribed by the following definitions. In order to stimulate such an organ it is necessary that it should receive simultaneously at least as many stimuli on its excitatory inputs as correspond to its threshold, and not a single stimulus on any one of its inhibitory inputs. If it has been thus stimulated, it will after a definite time delay (which is assumed to be always the same, and may be used to define the unit of time) emit an output pulse. This pulse can be taken by appropriate connections to any number of inputs of other neurons (also to any of its own inputs) and will produce at each of these the same type of input stimulus as the ones described above.

It is, of course, understood that this is an oversimplification of the actual functioning of a neuron. I have already discussed the character, the limitations, and the advantages of the axiomatic method. (See pages 2 and 10.) They all apply here, and the discussion which follows is to be taken in this sense.

McCulloch and Pitts have used these units to build up complicated networks which may be called "formal neural networks:" Such a system is built up of any number of these units, with their inputs and outputs suitably interconnected with arbitrary complexity. The "functioning" of such a network may be defined by singling out some of the inputs of the entire system and some of its outputs, and then describing what original stimuli on the former are to cause what ultimate stimuli on the latter.

The Main Result of the McCulloch-Pitts Theory. McCulloch and Pitts' important result is that any functioning in this sense which can be defined at all logically, strictly, and unambiguously in a finite number of words can also be realized by such a formal neural network.

It is well to pause at this point and to consider what the implications are. It has often been claimed that the activities and functions of the human nervous system are so complicated that no ordinary mechanism could possibly perform them. It has also been attempted to name specific functions which by their nature exhibit this limitation. It has been attempted to show that such specific functions, logically, completely described, are per se unable of mechanical, neural realization. The McCulloch-Pitts result puts an end to this. It proves that anything that can be exhaustively and unambiguously described, anything that can be completely and unambiguously put into words, is ipso facto realizable by a suitable finite neural network. Since the converse statement is obvious, we can therefore say that there is no difference between the possibility of describing a real or imagined mode of behavior completely and unambiguously in words, and the possibility of realizing it by a finite formal neural network. The two concepts are co-extensive. A difficulty of principle embodying any mode of behavior in such a network can exist only if we are also unable to describe that behavior completely.

Thus the remaining problems are these two. First, if a certain mode of behavior can be effected by a finite neural network, the question still remains whether that network can be realized within a practical size, specifically, whether it will fit into the physical limitations of the organism in question. Second, the question arises whether every existing mode of behavior can really be put completely and unambiguously into words. The first problem is, of course, the ultimate problem of nerve physiology, and I shall not attempt to go into it any further here. The second question is of a different character, and it has interesting logical connotations.

Interpretations of This Result. There is no doubt that any special phase of any conceivable form of behavior can be described "completely and unambiguously" in words. This description may be lengthy, but it is always possible. To deny it would amount to adhering to a form of logical mysticism which is surely far from most of us. It is, however, an important limitation, that this applies only to every element separately, and it is far from clear how it will apply to the entire syndrome of behavior. To be more specific, there is no difficulty in describing how an organism might be able to identify any two rectilinear triangles, which appear on the retina, as belonging to the same category "triangle." There is also no difficulty in adding to this, that numerous other objects, besides regularly drawn rectilinear triangles, will also be classified and identified as triangles—triangles whose sides are curved, triangles whose sides are not fully drawn, triangles that are indicated merely by a more or less homogeneous shading of their interior, etc. The more completely we attempt to describe everything that may conceivably fall under this heading, the longer the description becomes. We may have a vague and uncomfortable feeling that a complete catalogue along such lines would not only be exceedingly long, but also unavoidably indefinite at its boundaries. Nevertheless, this may be a possible operation.

All of this, however, constitutes only a small fragment of the more general concept of identification of analogous geometrical entities. This, in turn, is only a microscopic piece of the general concept of analogy. Nobody would attempt to describe and define within any practical amount of space the general concept of analogy which dominates our interpretation of vision. There is no basis for saying whether such an enterprise would require thousands or millions or altogether impractical numbers of volumes. Now it is perfectly possible that the simplest and only practical way actually to say what constitutes a visual analogy consists in giving a description of the connections of the visual brain. We are dealing here with parts of logics with which we have practically no past experience. The order of complexity is out of all proportion to anything we have ever known. We have no right to assume that the logical notations and procedures used in the past are suited to this part of the subject. It is not at all certain that in this domain a real object might not constitute the simplest description of itself, that is, any attempt to describe it by the usual literary or formal-logical method may lead to something less manageable and more involved. In fact, some results in modern logic would tend to indicate that phenomena like this have to be expected when we come to really complicated entities. It is, therefore, not at all unlikely that it is futile to look for a precise logical concept, that is, for a precise verbal description, of "visual analogy." It is possible that the connection pattern of the visual brain itself is the simplest logical expression or definition of this principle.

Obviously, there is on this level no more profit in the McCulloch-Pitts result. At this point it only furnishes another illustration of the situation outlined earlier. There is an equivalence between logical principles and their embodiment in a neural network, and while in the simpler cases the principles might furnish a simplified expression of the network, it is quite possible that in cases of extreme complexity the reverse is true. All of this does not alter my belief that a new, essentially logical, theory is called for in order to understand high-complication automata and, in particular, the central nervous system. It may be, however, that in this process logic will have to undergo a pseudo-morphosis to neurology to a much greater extent than the reverse. The foregoing analysis shows that one of the relevant things we can do at this moment with respect to the theory of the central nervous system is to point out the directions in which the real problem does not lie.

THE CONCEPT OF COMPLICATION; SELF-REPRODUCTION

The Concept of Complication. The discussions so far have shown that high complexity plays an important role in any theoretical effort relating to automata, and that this concept, in spite of its prima facie quantitative character, may in fact stand for something qualitative—for a matter of principle. For the remainder of my discussion I will consider a remoter implication of this concept, one which makes one of the qualitative aspects of its nature even more explicit. There is a very obvious trait, of the "vicious circle" type, in nature, the simplest expression of which is the fact that very complicated organisms can reproduce themselves.

We are all inclined to suspect in a vague way the existence of a concept of "complication." This concept and its putative properties have never been clearly formulated. We are, however, always tempted to assume that they will work in this way. When an automaton performs certain operations, they must be expected to be of a lower degree of complication than the automaton itself. In particular, if an automaton has the ability to construct another one, there must be a decrease in complication as we go from the parent to the construct. That is, if A can produce B, then A in some way must have contained a complete description of B. In order to make it effective, there must be, furthermore, various arrangements in A that see to it that this description is interpreted and that the constructive operations that it calls for are carried out. In this sense, it would therefore seem that a certain degenerating tendency must be expected, some decrease in complexity as one automaton makes another automaton.

Although this has some indefinite plausibility to it, it is in clear contradiction with the most obvious things that go on in nature. Organisms reproduce themselves, that is, they produce new organisms with no decrease in complexity. In addition, there are long periods of evolution during which the complexity is even increasing. Organisms are indirectly derived from others which had lower complexity.

Thus there exists an apparent conflict of plausibility and evidence, if nothing worse. In view of this, it seems worth while to try to see whether there is anything involved here which can be formulated rigorously.

So far I have been rather vague and confusing, and not unintentionally at that. It seems to me that it is otherwise impossible to give a fair impression of the situation that exists here. Let me now try to become specific.

Turing's Theory of Computing Automata. The English logician, Turing, about twelve years ago attacked the following problem.

He wanted to give a general definition of what is meant by a computing automaton. The formal definition came out as follows:

An automaton is a "black box," which will not be described in detail but is expected to have the following attributes. It possesses a finite number of states, which need be prima facie characterized only by stating their number, say n, and by enumerating them accordingly: 1, 2, . . . n. The essential operating characteristic of the automaton consists of describing how it is caused to change its state, that is, to go over from a state i into a state j. This change requires some interaction with the outside world, which will be standardized in the following manner. As far as the machine is concerned, let the whole outside world consist of a long paper tape. Let this tape be, say, 1 inch wide, and let it be subdivided into fields (squares) 1 inch long. On each field of this strip we may or may not put a sign, say, a dot, and it is assumed that it is possible to erase as well as to write in such a dot. A field marked with a dot will be called a "1," a field unmarked with a dot will be called a "0." (We might permit more ways of marking, but Turing showed that this is irrelevant and does not lead to any essential gain in generality.) In describing the position of the tape relative to the automaton it is assumed that one particular field of the tape is under direct inspection by the automaton, and that the automaton has the ability to move the tape forward and backward, say, by one field at a time. In specifying this, let the automaton be in the state i (= 1 . . . , n), and let it see on the tape an e (= 0, 1). It will then go over into the state j (= 0, 1, . . ., n) move the tape by p fields (p = 0, +1, -1; +1 is a move forward, -1 is a move backward), and inscribe into the new field that it sees f (= 0, 1; inscribing 0 means erasing; inscribing 1 means putting in a dot). Specifying j, p, f as functions of i, e is then the complete definition of the functioning of such an automaton.

Turing carried out a careful analysis of what mathematical processes can be effected by automata of this type. In this connection he proved various theorems concerning the classical "decision problem" of logic, but I shall not go into these matters here. He did, however, also introduce and analyze the concept of a "universal automaton," and this is part of the subject that is relevant in the present context.

An infinite sequence of digits e (= 0, 1) is one of the basic entities in mathematics. Viewed as a binary expansion, it is essentially equivalent to the concept of a real number. Turing, therefore, based his consideration on these sequences.

He investigated the question as to which automata were able to construct which sequences. That is, given a definite law for the formation of such a sequence, he inquired as to which automata can be used to form the sequence based on that law. The process of "forming" a sequence is interpreted in this manner. An automaton is able to "form" a certain sequence if it is possible to specify a finite length of tape, appropriately marked, so that, if this tape is fed to the automaton in question, the automaton will thereupon write the sequence on the remaining (infinite) free portion of the tape. This process of writing the infinite sequence is, of course, an indefinitely continuing one. What is meant is that the automaton will keep running indefinitely and, given a sufficiently long time, will have inscribed any desired (but of course finite) part of the (infinite) sequence. The finite, premarked, piece of tape constitutes the "instruction" of the automaton for this problem.

An automaton is "universal" if any sequence that can be produced by any automaton at all can also be solved by this particular automaton. It will, of course, require in general a different instruction for this purpose.

The Main Result of the Turing Theory. We might expect a priori that this is impossible. How can there be an automaton which is at least as effective as any conceivable automaton, including, for example, one of twice its size and complexity?

Turing, nevertheless, proved that this is possible. While his construction is rather involved, the underlying principle is nevertheless quite simple. Turing observed that a completely general description of any conceivable automaton can be (in the sense of the foregoing definition) given in a finite number of words. This description will contain certain empty passages—those referring to the functions mentioned earlier (j, p, f in terms of i, e), which specify the actual functioning of the automaton. When these empty passages are filled in, we deal with a specific automaton. As long as they are left empty, this schema represents the general definition of the general automaton. Now it becomes possible to describe an automaton which has the ability to interpret such a definition. In other words, which, when fed the functions that in the sense described above define a specific automaton, will thereupon function like the object described. The ability to do this is no more mysterious than the ability to read a dictionary and a grammar and to follow their instructions about the uses and principles of combinations of words. This automaton, which is constructed to read a description and to imitate the object described, is then the universal automaton in the sense of Turing. To make it duplicate any operation that any other automaton can perform, it suffices to furnish it with a description of the automaton in question and, in addition, with the instructions which that device would have required. for the operation under consideration.

Broadening of the Program to Deal with Automata That Produce Automata. For the question which concerns me here, that of "self-reproduction" of automata, Turing's procedure is too narrow in one respect only. His automata are purely computing machines. Their output is a piece of tape with zeros and ones on it. What is needed for the construction to which I referred is an automaton whose output is other automata. There is, however, no difficulty in principle in dealing with this broader concept and in deriving from it the equivalent of Turing's result.

The Basic Definitions. As in the previous instance, it is again of primary importance to give a rigorous definition of what constitutes an automaton for the purpose of the investigation. First of all, we have to draw up a complete list of the elementary parts to be used. This list must contain not only a complete enumeration but also a complete operational definition of each elementary part. It is relatively easy to draw up such a list, that is, to write a catalogue of "machine parts" which is sufficiently inclusive to permit the construction of the wide variety of mechanisms here required, and which has the axiomatic rigor that is needed for this kind of consideration. The list need not be very long either. It can, of course, be made either arbitrarily long or arbitrarily short. It may be lengthened by including in it, as elementary parts, things which could be achieved by combinations of others. It can be made short—in fact, it can be made to consist of a single unit—by endowing each elementary part with a multiplicity of attributes and functions. Any statement on the number of elementary parts required will therefore represent a common-sense compromise, in which nothing too complicated is expected from any one elementary part, and no elementary part is made to perform several, obviously separate, functions. In this sense, it can be shown that about a dozen elementary parts suffice. The problem of self-reproduction can then be stated like this: Can one build an aggregate out of such elements in such a manner that if it is put into a reservoir, in which there float all these elements in large numbers, it will then begin to construct other aggregates, each of which will at the end turn out to be another automaton exactly like the original one? This is feasible, and the principle on which it can be based is closely related to Turing's principle outlined earlier.

Outline of the Derivation of the Theorem Regarding Self-reproduction. First of all, it is possible to give a complete description of every thing that is an automaton in the sense considered here. This description is to be conceived as a general one, that is, it will again contain empty spaces. These empty spaces have to be filled in with the functions which describe the actual structure of an automaton. As before, the difference between these spaces filled and unfilled is the difference between the description of a specific automaton and the general description of a general automaton. There is no difficulty of principle in describing the following automata.

(a) Automaton A, which when furnished the description of any other automaton in terms of appropriate functions, will construct that entity. The description should in this case not be given in the form of a marked tape, as in Turing's case, because we will not normally choose a tape as a structural element. It is quite easy, however, to describe combinations of structural elements which have all the notational properties of a tape with fields that can be marked. A description in this sense will be called an instruction and denoted by a letter I.

"Constructing" is to be understood in the same sense as before. The constructing automaton is supposed to be placed in a reservoir in which all elementary components in large numbers are floating, and it will effect its construction in that milieu. One need not worry about how a fixed automaton of this sort can produce others which are larger and more complex than itself. In this case the greater size and the higher complexity of the object to be constructed will be reflected in a presumably still greater size of the instructions I that have to be furnished. These instructions, as pointed out, will have to be aggregates of elementary parts. In this sense, certainly, an entity will enter the process whose size and complexity is determined by the size and complexity of the object to be constructed.

In what follows, all automata for whose construction the facility A will be used are going to share with A this property. All of them will have a place for an instruction I, that is, a place where such an instruction can be inserted. When such an automaton is being described (as, for example, by an appropriate instruction), the specification of the location for the insertion of an instruction I in the foregoing sense is understood to form a part of the description. We may, therefore, talk of "inserting a given instruction I into a given automaton," without any further explanation.

(b) Automaton B, which can make a copy of any instruction I that is furnished to it. 1 is an aggregate of elementary parts in the sense outlined in (a), replacing a tape. This facility will be used when 1 furnishes a description of another automaton. In other words, this automaton is nothing more subtle than a "reproducer"—the machine which can read a punched tape and produce a second punched tape that is identical with the first. Note that this automaton, too, can produce objects which are larger and more complicated than itself. Note again that there is nothing surprising about it. Since it can only copy, an object of the exact size and complexity of the output will have to be furnished to it as input.

After these preliminaries, we can proceed to the decisive step.

(c) Combine the automata A and B with each other, and with a control mechanism C which does the following. Let A be furnished with an instruction I (again in the sense of [a] and [b]), then C will first cause A to construct the automaton which is described by this instruction I. Next C will cause B to copy the instruction I referred to above, and insert the copy into the automaton referred to above, which has just been constructed by A. Finally, C will separate this construction from the system A + B + C and "turn it loose" as an independent entity.

(d) Denote the total aggregate A + B + C by D.

(e) In order to function, the aggregate D = A + B + C must be furnished with an instruction I, as described above. This instruction, as pointed out above, has to be inserted into A. Now form an instruction I_D which describes this automaton D, and insert I_D into A within D. Call the aggregate which now results E.

E is clearly self-reproductive. Note that no vicious circle is involved. The decisive step occurs in E, when the instruction I_D, describing D, is constructed and attached to D. When the construction (the copying) of I_D called for, D exists already, and it is in no wise modified by the construction of I_D. I_D is simply added to form E. Thus there is a definite chronological and logical order in which D and I_D have to be formed, and the process is legitimate and proper according to the rules of logic.

Interpretations of This Result and of Its Immediate Extensions. The description of this automaton E has some further attractive sides, into which I shall not go at this time at any length. For instance, it is quite clear that the instruction I is roughly effecting the functions of a gene. It is also clear that the copying mechanism B performs the fundamental act of reproduction, the duplication of the genetic material, which is clearly the fundamental operation in the multiplication of living cells. It is also easy to see how arbitrary alterations of the system E, and in particular of I_D can exhibit certain typical traits which appear in connection with mutation, lethally as a rule, but with a possibility of continuing reproduction with a modification of traits. It is, of course, equally clear at which point the analogy ceases to be valid. The natural gene does probably not contain a complete description of the object whose construction its presence stimulates. It probably contains only general pointers, general cues. In the generality in which the foregoing consideration is moving, this simplification is not attempted. It is, nevertheless, clear that this simplification, and others similar to it, are in themselves of great and qualitative importance. We are very far from any real understanding of the natural processes if we do not attempt to penetrate such simplifying principles.

Small variations of the foregoing scheme also permit us to construct automata which can reproduce themselves and, in addition, construct others. (Such an automaton performs more specifically what is probably a—if not the—typical gene function, self-reproduction plus production—or stimulation of production—of certain specific enzymes.) Indeed, it suffices to replace the I_D by an instruction I_{D+F}, which describes the automaton D plus another given automaton F. Let D, with I_{D+F} inserted into A within it, be designated by E_F. This E_F clearly has the property already described. It will reproduce itself, and, besides, construct F.

Note that a "mutation" of E in E_F, which takes place within the F-part of I_{D+F} in E_F, is not lethal. If it replaces F by F', it changes E_F into $E_{F'}$ that is, the "mutant" is still self-reproductive; but its by-product is changed—F' instead of F. This is, of course, the typical non-lethal mutant.

All these are very crude steps in the direction of a systematic theory of automata. They represent, in addition, only one particular direction. This is, as I indicated before, the direction towards forming a rigorous concept of what constitutes "complication." They illustrate that "complication" on its lower levels is probably degenerative, that is, that every automaton that can produce other automata will only be able to produce less complicated ones. There is, however, a certain minimum level where this degenerative characteristic ceases to be universal. At this point automata which can reproduce themselves, or even construct higher entities, become possible. This fact, that complication, as well as organization, below a certain minimum level is degenerative, and beyond that level can become self-supporting and even increasing, will clearly play an important role in any future theory of the subject.

DISCUSSION

Dr MC CULLOCH: I confess that there is nothing I envy Dr. von Neumann more than the fact that the machines with which he has to cope are those for which he has, from the beginning, a blueprint of what the machine is supposed to do and how it is supposed to do it. Unfortunately for us in the biological sciences—or, at least, in psychiatry—we are presented with an alien, or enemy's, machine. We do not know exactly what the machine is supposed to do and certainly we have no blueprint of it. In attacking our problems, we only know, in psychiatry, that the machine is producing wrong answers. We know that, because of the damage by the machine to the machine itself and by its running amuck in the world. However, what sort of difficulty exists in that machine is no easy matter to determine.

As I see it what we need first and foremost is not a correct theory, but some theory to start from, whereby we may hope to ask a question so that we'll get an answer, if only to the effect that our notion was entirely erroneous. Most of the time we never even get around to asking the question in such a form that it can have an answer.

I'd like to say, historically, how I came to be interested in this particular problem, if you'll forgive me, because it does bear on this matter. I came, from a major interest in philosophy and mathematics, into psychology with the problem of how a thing like mathematics could ever arise-what sort of a thing it was. For that reason, I gradually shifted into psychology and thence, for the reason that I again and again failed to find the significant variables, I was forced into neurophysiology. The attempt to construct a theory in a field like this, so that it can be put to any verification, is tough. Humorously enough, I started entirely at the wrong angle, about 1919, trying to construct a logic for transitive verbs. That turned out to be as mean a problem as modal logic, and it was not until I saw Turing's paper that I began to get going the right way around, and with Pitts' help formulated the required logical calculus. What we thought we were doing (and I think we succeeded fairly well) was treating the brain as a Turing machine; that is, as a device which could perform the kind of functions which a brain must perform if it is only to go wrong and have a psychosis. The important thing was, for us, that we had to take a logic and subscript it for the time of the occurrence of a signal (which is, if you will, no more than a proposition on the move). This was needed in order to construct theory enough to be able to state how a nervous system could do anything. The delightful thing is that the very simplest set of appropriate assumptions is sufficient to show that a nervous system can compute any computable number. It is that kind of a device, if you like—a Turing machine.

The question at once arose as to how it did certain of the things that it did do. None of the theories tell you how a particular operation is carried out, any more than they tell you in what kind of a nervous system it is carried out, or any more than they tell you in what part of a computing machine it is carried out. For that you have to have the wiring diagram or the prescription for the relations of the gears.

This means that you are compelled to study anatomy, and to require of the anatomist the things he has rarely given us in sufficient detail. I taught neuro-anatomy while I was in medical school, but until the last year or two I have not been in a position to ask any neuroanatomist for the precise detail of any structure. I had no physiological excuse for wanting that kind of information. Now we are beginning to need it.

Dr. GERARD: I have had the privilege of hearing Dr. von Neumann speak on various occasions, and I always find myself in the delightful but difficult role of hanging on to the tail of a kite. While I can follow him, I can't do much creative thinking as we go along. I would like to ask one question, though, and suspect that it may be in the minds of others. You have carefully stated, at several points in your discourse, that anything that could be put into verbal form—into a question with words—could be solved. Is there any catch in this? What is the implication of just that limitation or. the question?

Dr. VON NEUMANN: I will try to answer, but my answer will have to be rather incomplete.

The first task that arises in dealing with any problem—more specifically, with any function of the central nervous system—is to formulate it unambiguously, to put it into words, in a rigorous sense. If a very complicated system—like the central nervous system—is involved, there arises the additional task of doing this "formulating," this "putting into words," with a number of words within reasonable limits—for example, that can be read in a lifetime. This is the place where the real difficulty lies.

In other words, I think that it is quite likely that one may give a purely descriptive account of the outwardly visible functions of the central nervous system in a humanly possible time. This may be 10 or 20 years—which is long, but not prohibitively long. Then, on the basis of the results of McCulloch and Pitts, one could draw within plausible time limitations a fictitious "nervous network" that can carry out all these functions. I suspect, however, that it will turn out to be much larger than the one that we actually possess. It is possible that it will prove to be too large to fit into the physical universe. What then? Haven't we lost the true problem in the process?

Thus the problem might better be viewed, not as one of imitating the functions of the central nervous system with just any kind of network, but rather as one of doing this with a network that will fit into the actual volume of the human brain. Or, better still, with one that can be kept going with our actual metabolistic "power supply" facilities, and that can be set up and organized by our actual genetic control facilities.

To sum up, I think that the first phase of our problem—the purely formalistic one, that one of finding any "equivalent network" at all—has been overcome by McCulloch and Pitts. I also think that much of the "malaise" felt in connection with attempts to "explain" the central nervous system belongs to this phase—and should therefore be considered removed. There remains, however, plenty of malaise due to the next phase of the problem, that one of finding an "equivalent network" of possible, or even plausible, dimensions and (metabolistic and genetic) requirements.

The problem, then, is not this: How does the central nervous system effect any one, particular thing? It is rather: How does it do all the things that it can do, in their full complexity? What are the principles of its organization? How does it avoid really serious, that is, lethal, malfunctions over periods that seem to average many decades?

Dr. GERARD: Did you mean to imply that there are unformulated problems?

Dr. VON NEUMANN: There may be problems which cannot be formulated with our present logical techniques.

Dr. WEISS: I take it that we are discussing only a conceivable and logically consistent, but not necessarily real, mechanism of the nervous system. Any theory of the real nervous system, however, must explain the facts of regulation—that the mechanism will turn out the same or an essentially similar product even after the network of pathways has been altered in many unpredictable ways. According to von Neumann, a machine can be constructed so as to contain safeguards against errors and provision for correcting errors when they occur. In this case the future contingencies have been taken into account in constructing the machine. In the case of the nervous system, evolution would have had to build in the necessary corrective devices. Since the number of actual interferences and deviations produced by natural variation and by experimenting neurophysiologists is very great, I question whether a mechanism in which all these innumerable contingencies have been foreseen, and the corresponding corrective measures build in, is actually conceivable.

Dr. VON NEUMANN: I will not try, of course, to answer the question as to how evolution came to any given point. I am going to make, however, a few remarks about the much more limited question regarding errors, foreseeing errors, and recognizing and correcting errors. An artificial machine may well be provided with organs which recognize and correct errors automatically. In fact, almost every well-planned machine contains some organs whose function is to do just this—always within certain limited areas. Furthermore, if any particular machine is given, it is always possible to construct a second machine which "watches" the first one, and which senses and possibly even corrects its errors. The trouble is, however, that now the second machine's errors are unchecked, that is, *quis custodiet ipsos custodes?* Building a third, a fourth, etc., machine for second order, third order, etc., checking merely shifts the problem. In addition, the primary and the secondary machine will, together, make more errors than the first one alone, since they have more components.

Some such procedure on a more modest scale may nevertheless make sense. One might know, from statistical experience with a certain machine or class of machines, which ones of its components malfunction most frequently, and one may then "supervise" these only, etc. Another possible approach, which permits a more general quantitative evaluation,. is this: Assume that one had a machine which has a probability of 10^{-10} to malfunction on any single operation, that is, which will, on the average, make one error for any 10^{10} operations. Assume that this machine has to solve a problem that requires 10^{12} operations. Its normal "unsupervised" functioning will, therefore, on the average, give 100 errors in a single problem, that is, it will be completely unusable.

Connect now three such machines in such a manner that they always compare their results after every single operation, and then proceed as follows. (a) If all three have the same result, they continue unchecked. (b) If any two agree with each other, but not with the third, then all three continue with the value agreed on by the majority. (c) If no two agree with each other, then all three stop.

This system will produce a correct result, unless at some point in the problem two of the three machines err simultaneously. The probability of two given machines erring simultaneously on a given operation is $10^{-10} \times 10^{-10} = 10^{-20}$. The probability of any two doing this on a given operation is 3×10^{-20} (there are three possible pairs to be formed among three individuals [machines]). The probability of this happening at all (that is, anywhere) in the entire problem is $10^{12} \times 3 \times 10^{-20} = 3 \times 10^{-8}$, about one in 33 million.

Thus there is only one chance in 33 million that this triad of machines will fail to solve the problem correctly—although each member of the triad alone had hardly any chance to solve it correctly.

Note that this triad, as well as any other conceivable automatic contraption, no matter how sophisticatedly supervised, still offers a logical possibility of resulting error—although, of course, only with a low probability. But the incidence (that is, the probability) of error has been significantly lowered, and this is all that is intended.

Dr. WEISS: In order to crystallize the issue, I want to reiterate that if you know the common types of errors that will occur in a particular machine, you can make provisions for the correction of these errors in constructing the machine. One of the major features of the nervous system, however, is its apparent ability to remedy situations that could not possibly have been foreseen. (The number of artificial interferences with the various apparatuses of the nervous system that can be applied without impairing the biologically useful response of the organism is infinite.) The concept of a nervous automaton should, therefore, not only be able to account for the normal operation of the nervous system but also for its relative stability under all kinds of abnormal situations.

Dr. VON NEUMANN: I do not agree with this conclusion. The argumentation that you have used is risky, and requires great care.

One can in fact guard against errors that are not specifically foreseen. These are some examples that show what I mean.

One can design and build an electrical automaton which will function as long as every resistor in it deviates no more than 10 per cent from its standard design value. You may now try to disturb this machine by experimental treatments which will alter its resistor values (as, for example, by heating certain regions in the machine). As long as no resistor shifts by more than 10 per cent, the machine will function right—no matter how involved, how sophisticated, how "unforeseen" the disturbing experiment is.

Or—another example—one may develop an armor plate which will resist impacts up to a certain strength. If you now test it, it will stand up successfully in this test, as long as its strength limit is not exceeded, no matter how novel the design of the gun, propellant, and projectile used in testing, etc.

It is clear how these examples can be transposed to neural and genetic situations.

To sum up: Errors and sources of errors need only be foreseen generically, that is, by some decisive traits, and not specifically, that is, in complete detail. And these generic coverages may cover vast territories, full of unforeseen and unsuspected—but, in fine, irrelevant—details.

Dr. Mc CULLOCH: How about designing computing machines so that if they were damaged in air raids, or what not, they could replace parts, or service themselves, and continue to work?

Dr. VON NEUMANN: These are really quantitative rather than qualitative questions. There is no doubt that one can design machines which, under suitable conditions, will repair themselves. A practical discussion is, however, rendered difficult by what I believe to be a rather accidental circumstance. This is, that we seem to be operating with much more unstable materials than nature does. A metal may seem to be more stable than a tissue, but, if a tissue is injured, it has a tendency to restore itself, while our industrial materials do not have this tendency, or have it to a considerably lesser degree. I don't think, however, that any question of principle is involved at this point. This reflects merely the present, imperfect state of our technology—a state that will presumably improve with time.

Dr. LASHLEY: I'm not sure that I have followed exactly the meaning of "error" in this discussion, but it seems to me the question of precision of the organic machine has been somewhat exaggerated. In the computing machines, the one thing we demand is precision; on the other hand, when we study the organism, one thing which we never find is accuracy or precision. In any organic reaction there is a normal, or nearly normal, distribution of errors around a mean. The mechanisms of reaction are statistical in character and their accuracy is only that of a probability distribution in the activity of enormous numbers of elements. In this respect the organism resembles the analogical rather than the digital machine. The invention of symbols and the use of memorized number series convert the organism into a digital machine, but the increase in accuracy is acquired at the sacrifice of speed. One can estimate the number of books on a shelf at a glance, with some error. To count them requires much greater time. As a digital machine the organism is inefficient. That is why you build computing machines.

Dr. VON NEUMANN*:* I would like to discuss this question of precision in some detail.

It is perfectly true that in all mathematical problems the answer is required with absolute rigor, with absolute reliability. This may, but need not, mean that it is also required with absolute precision. In most problems for the sake of which computing machines are being built—mostly problems in various parts of applied mathematics, mathematical physics—the precision that is wanted is quite limited. That is, the data of the problem are only given to limited precision, and the result is only wanted to limited precision. This is quite compatible with absolute mathematical rigor, if the sensitivity of the result to changes in the data as well as the limits of uncertainty (that is, the amount of precision) of the result for given data are (rigorously) known.

The (input) data in physical problems are often not known to better than a few (say 5) per cent. The result may be satisfactory to even less precision (say 10 per cent). In this respect, therefore, the difference of outward precision requirements for an (artificial) computing machine and a (natural) organism need not at all be decisive. It is merely quantitative, and the quantitative factors involved need not be large at that.

The need for high previsions in the internal functioning of (artificial) computing machines is due to entirely different causes—and these may well be operating in (natural) organisms too. By this I do not mean that the arguments that follow should be carried over too literally to organisms. In fact, the "digital method" used in computing may be entirely alien to the nervous system. The discrete pulses used in neural communications look indeed more like "counting" by numeration than like a "digitalization." (In many cases, of course, they may express a logical code—this is quite similar to what goes on in computing machines.) I will, nevertheless, discuss the specifically "digital" procedure of our computing machine, in order to illustrate how subtle the distinction between "external" and "internal" precision requirements can be.

In a computing machine numbers may have to be dealt with as aggregates of 10 or more decimal places. Thus an internal precision of one in 10 billion or more may be needed, although the data are only good to one part in 20 (5 per cent), and the result is only wanted to one part in 10 (10 per cent). The reason for this strange discrepancy is that a fast machine will only be used on long and complicated problems. Problems involving 100 million multiplications will not be rarities. In a 4-decimal-place machine every multiplication introduces a "round-off" error of one part in 10,000; in a 6-place machine this is one part in a million; in a 10-place machine it is one part in 10 billion. In a problem of the size indicated above, such errors will occur 100 million times. They will be randomly distributed, and it follows therefore from the rules of mathematical statistics that the total error will probably not be 100 million times the individual (round-off) error, but about the square root of 100 million times, that is, about 10,000 times. A precision of 10 per cent—one part in 10—in the result should therefore require 10,000 times more precision than this on individual steps (multiplication round-offs): namely, one part in 100,000, that is, 5 decimal places. Actually, more will be required because the (round-off) errors made in the earlier parts of the calculation are frequently "amplified" by the operations of the subsequent parts of the calculation. For these reasons 8 to 10 decimal places are probably a minimum for such a machine, and actually many large problems may well require more.

Most analogy computing machines have much less precision than this (on elementary operations). The electrical ones usually one part in 100 or 1000, the best mechanical ones (the most advanced "differential analyzers") one part in 10,000 or 50,000. The virtue of the digital method is that it will, with componentry of very limited precision, give almost any precision on elementary operations. If one part in a million is wanted, one will use 6 decimal digits; if one part in 10 billions is wanted, one need only increase the number of decimal digits to 10; etc. And yet the individual components need only be able to distinguish reliably 10 different states (the 10 decimal digits from 0 to 9), and by some simple logical and organizational tricks one can even get along with components that can only distinguish two states!

I suspect that the central nervous system, because of the great complexity of its tasks, also faces problems of "internal" precision or reliability. The all-or-none character of nervous impulses may be connected with some technique that meets this difficulty, and this—unknown—technique might well be related to the digital system that we use in computing, although it is probably very different from the digital system in its technical details. We seem to have no idea as to what this technique is. This is again an indication of how little we know. I think, however, that the digital system of computing is the only thing known to us that holds any hope of an even remote affinity with that unknown, and merely postulated, technique.

Dr. MCCULLOCH: I want to make a remark in partial answer to Dr. Lashley. I think that the major woe that I have always encountered in considering the behavior of organisms was not in such procedures as hitting a bull's-eye or judging a distance, but in mathematics and logic. After all, Vega did compute log tables to thirteen places. He made some four hundred and thirty errors, but the total precision of the work of that organism is simply incredible to me.

Dr. LASHLEY: You must keep in mind that such an achievement is not the product of a single elaborate integration but represents a great number of separate processes which are, individually, simple discriminations far above threshold values and which do not require great accuracy of neural activity.

Dr. HALSTEAD: As I listened to Dr. von Neumann's beautiful analysis of digital and analogous devices, I was impressed by the conceptual parsimony with which such systems may be described. We in the field of organic behavior are not yet so fortunate. Our parsimonies, for the most part, are still to be attained. There is virtually no class of behaviors which can at present be described with comparable precision. Whether such domains as thinking, intelligence, learning, emoting, language, perception, and response represent distinctive processes or only different attitudinal sets of the organism is by no means clear. It is perhaps for this reason that Dr. von Neumann did not specify the class or classes of behaviors which his automata simulate.

As Craik pointed out several years ago,[*] it isn't quite logically air-tight to compare the operations of models with highly specified ends with organic behaviors only loosely specified either hierarchically or as to ends. Craik's criterion was that our models must bear a proper "relation structure" to the steps in the processes simulated. The rules of the game are violated when we introduce gremlins, either good or bad gremlins, as intervening variables. It is not clear to me whether von Neumann means "noise" as a good or as a bad gremlin. I presume it is a bad one when it is desired to maximize "rationality" in the outcome. It is probable that rationality characterizes a restricted class of human behavior. I shall later present experimental evidence that the same normal or brain-injured man also produces a less restricted class of behavior which is "arational" if not irrational. I suspect that von Neumann biases his automata towards rationality by careful regulation

[*] Nature of Explanation, London, Cambridge University Press, 1943.

of the energies of the substrate. Perhaps he would gain in similitude, however, were he to build unstable power supplies into his computers and observe the results.

It seems to me that von Neumann is approximating in his computers some of the necessary operations in the organic process recognized by psychologists under the term "abstraction." Analysis of this process of ordering to a criterion in brain-injured individuals suggests that three classes of outcome occur. First, there is the pure category (or "universal"); second, there is the partial category; and third, there is the phenomenal or non-recurrent organization. Operationalism restricts our concern to the first two classes. However, these define the third. It is probably significant that psychologists such as Spearman and Thurstone have made considerable progress in describing these outcomes in mathematical notation.

Dr. LORENTE DE NÓ: I began my training in a very different manner from Dr. McCulloch. I began as an anatomist and became interested in physiology much later. Therefore, I am still very much of an anatomist, and visualize everything in anatomical terms. According to your discussion, Dr. von Neumann, of the McCulloch and Pitts automaton, anything that can be expressed in words can be performed by the automaton. To this I would say that I can remember what you said, but that the McCulloch-Pitts automaton could not remember what you said. No, the automaton does not function in the way that our nervous system does, because the only way in which that could happen, as far as I can visualize, is by having some change continuously maintained. Possibly the automaton can be made to maintain memory, but the automaton that does would not have the properties of our nervous system. We agree on that, I believe. The only thing that I wanted was to make the fact clear.

Dr. VON NEUMANN: One of the peculiar features of the situation, of course, is that you can make a memory out of switching organs, but there are strong indications that this is not done in nature. And, by the way, it is not very efficient, as closer analysis shows.